



Published in final edited form as:

Lang Cogn Process. 2011 January 1; 26(7): 952–981. doi:10.1080/01690960903498424.

The DIVA model: A neural theory of speech acquisition and production

Jason A. Tourville¹ and Frank H. Guenther^{1,2,3}

¹Department of Cognitive and Neural Systems, Boston University, 677 Beacon Street, Boston, MA, 02215, Telephone: (617) 353-5765, Fax Number: (617) 353-7755, guenther@cns.bu.edu

²Division of Health Sciences and Technology, Harvard University - Massachusetts Institute of Technology

³Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital

Abstract

The DIVA model of speech production provides a computationally and neuroanatomically explicit account of the network of brain regions involved in speech acquisition and production. An overview of the model is provided along with descriptions of the computations performed in the different brain regions represented in the model. The latest version of the model, which contains a new right-lateralized *feedback control map* in ventral premotor cortex, will be described, and experimental results that motivated this new model component will be discussed. Application of the model to the study and treatment of communication disorders will also be briefly described.

INTRODUCTION

With the proliferation of functional brain imaging studies, a consensus regarding the brain areas underlying speech motor control is building (e.g. Indefrey & Levelt, 2004; Turkeltaub, Eden, Jones & Zeffiro, 2002). Bohland and Guenther (2006) have described a “minimal network” of brain regions involved in speech production that includes bilateral medial and lateral frontal cortex, parietal cortex, superior temporal cortex, the thalamus, basal ganglia, and the cerebellum. It is little surprise that these regions are commonly associated with the planning and execution of movements (primary sensorimotor and premotor cortex, the supplementary motor area, the cerebellum, thalamus, and basal ganglia) and those associated with acoustic and phonological processing of speech sounds (the superior temporal gyrus). A complete, mechanistic account of the role played by each region during speech production and how they interact to produce fluent speech is still lacking.

The goal of our research program over the past sixteen years has been to improve our understanding of the neural mechanisms that underlie speech motor control. Over that time we have developed a computational model of speech acquisition and production called the DIVA model (Guenther, 1994; Guenther, 1995; Guenther, Ghosh & Tourville, 2006; Guenther, Hampson & Johnson, 1998). DIVA is an adaptive neural network that describes the sensorimotor interactions involved in articulator control during speech production. The model has been used to guide a number of behavioral and functional imaging studies of speech processing (e.g., Bohland & Guenther, 2006; Ghosh, Tourville & Guenther, 2008; Guenther, Espy-Wilson, Boyce, Matthies, Zandipour et al., 1999; Lane, Denny, Guenther,

Hanson, Marrone et al., 2007; Lane, Denny, Guenther, Matthies, Menard et al., 2005; Lane, Matthies, Guenther, Denny, Perkell et al., 2007; Nieto-Castanon, Guenther, Perkell & Curtin, 2005; Perkell, Guenther, Lane, Matthies, Stockmann et al., 2004; Perkell, Matthies, Tiede, Lane, Zandipour et al., 2004; Tourville, Reilly & Guenther, 2008). The mathematically explicit nature of the model allows for straightforward comparisons of hypotheses generated from simulations of experimental conditions to empirical data. Simulations of the model generate predictions regarding the expected acoustic (e.g., formant frequencies), somatosensory (e.g., articulator positions), learning rates, and activity levels within specific model components. Experiments are designed to test these predictions, and the empirical findings are, in turn, used to further refine the model.

In its current form, the DIVA model provides a unified explanation of a number of speech production phenomena including motor equivalence (variable articulator configurations that produce the same acoustic output), contextual variability, anticipatory and carryover coarticulation, velocity/distance relationships, speaking rate effects, and speaking skill acquisition and retention throughout development (e.g., Callan, Kent, Guenther & Vorperian, 2000; Guenther, 1994; Guenther, 1995; Guenther et al., 2006; Guenther et al., 1998; Nieto-Castanon et al., 2005). Because it can account for such a wide array of data, the DIVA model has provided the theoretical framework for a number of investigations of normal and disordered speech production. Predictions from the model have guided studies of the role of auditory feedback in normally hearing persons, deaf persons, and persons who have recently regained some hearing through the use of cochlear implants (Lane et al., 2007; Perkell, Denny, Lane, Guenther, Matthies et al., 2007; Perkell, Guenther, Lane, Matthies, Perrier et al., 2000; Perkell et al., 2004; Perkell et al., 2004). The model has also been employed in investigations of the etiology of stuttering (Max, Guenther, Gracco, Ghosh & Wallace, 2004), and acquired apraxia of speech (Robin, Guenther, Narayana, Jacks, Tourville et al., 2008; Terband, Maassen, Brumberg & Guenther, 2008).

In this review, the key concepts of the DIVA model are described with a focus on recent modifications to the model. Our investigations of the brain regions involved in feedback-based articulator control have motivated the addition of a lateralized *feedback control map* in ventral premotor cortex of the right hemisphere. Additional brain regions known to contribute to speech motor control have also been incorporated into the model. Projections originating in the supplementary motor area and passing through the basal ganglia and thalamus are hypothesized to serve as gates on the outflow of motor commands. Support for these modifications and the impact they have on the model are discussed below.

DIVA MODEL OVERVIEW

The DIVA model, schematized in Figure 1, consists of integrated feedforward and feedback control subsystems. Together, they learn to control a simulated vocal tract, a modified version of the synthesizer described by Maeda (1990). Once trained, the model takes a speech sound as input, and generates a time varying sequence of articulator positions that command movements of the simulated vocal tract that produce the desired sound. Each block in Figure 1 corresponds to a set of neurons that constitute a *neural representation*. When describing the model, the term *map* is used to refer to such a set of cells, represented by boxes in Figure 1. The term *mapping* is used to refer to a transformation from one neural representation to another. These transformations are represented by arrows in Figure 1 and are assumed to be carried out by filtering cell activations in one map through synapses projecting to another map. The synaptic weights are learned during a babbling phase meant to coarsely represent that typically experienced by a normally developing infant (e.g., Oller & Eilers, 1988). Random movements of the speech articulators provide tactile, proprioceptive, and auditory feedback signals that are used to learn the mappings between

the different neural representations. After babbling, the model can quickly learn to produce new sounds from audio samples provided to it.

An important part of the model's development has been assigning components of the model to corresponding regions of the brain (see anatomical labels in boxes Figure 1). Model components were mapped to locations in the Montreal Neurological Institute (MNI; Mazziotta, Toga, Evans, Fox, Lancaster et al., 2001) standard reference frame based on relevant neuroanatomical and neurophysiological studies (Guenther et al., 2006). The assignment of component locations is based on the synthesis of a large body of behavioral, neurophysiological, lesion, and neuroanatomical data. A majority of these data were derived from studies focused specifically on speech processes; however, studies of other modalities (e.g., non-orofacial motor control) also contributed. Associating model components to brain regions allows i) the generation of neuroanatomically specified hypotheses regarding the neural processes underlying speech motor control from a unified theoretical framework, and ii) a comparison of model dynamics with past, present, and future clinical and physiological findings regarding the functional neuroanatomy of speech processes.

Blood oxygenation level-dependent (BOLD) functional magnetic resonance imaging (fMRI; Belliveau, Kwong, Kennedy, Baker, Stern et al., 1992; Kwong, Belliveau, Chesler, Goldberg, Weisskoff et al., 1992; Ogawa, Menon, Tank, Kim, Merkle et al., 1993) has provided a powerful tool for the non-invasive study of human brain function. BOLD signal provides an indirect measure of neural activity. The coupling between neural activity and changes in local blood oxygen levels remains a matter of debate. While counter arguments have been made (Mukamel, Gelbard, Arieli, Hasson, Fried et al., 2005), consensus is building, around the hypotheses that BOLD signal is correlated with local field potentials (Goense & Logothetis, 2008; Logothetis, Pauls, Augath, Trinath & Oeltermann, 2001; Mathiesen, Caesar, Akgoren & Lauritzen, 1998; Viswanathan & Freeman, 2007). Local field potentials are thought to reflect local synaptic activity, i.e., a weighted sum of the inputs to a given region (Raichle & Mintun, 2006).

The development of BOLD fMRI has proven particularly beneficial to the study of speech given its uniquely human nature. The past decade and a half of imaging research has provided a tremendous amount of functional data regarding the brain regions involved in both speech (e.g., Fiez & Petersen, 1998; Ghosh, Bohland & Guenther, 2003; Indefrey & Levelt, 2004; Riecker, Ackermann, Wildgruber, Meyer, Dogil et al., 2000; Soros, Sokoloff, Bose, McIntosh, Graham et al., 2006; Turkeltaub et al., 2002), including the identification of a "minimal network" of the brain regions involved in speech production (Bohland & Guenther, 2006). A clear picture of the contributions made by each region, and how these regions interact during speech production remains elusive, however.

We believe that the continued study of the neural mechanisms of speech will benefit from the combined use of functional neuroimaging and computational modeling. For this purpose, the proposed locations of model components have been mapped to anatomical landmarks of the canonical brain provided with the SPM image analysis software package (Friston, Holmes, Poline, Grasby, Williams et al., 1995; <http://www.fil.ion.ucl.ac.uk/spm/>). The SPM canonical brain is a popular substrate for presenting neuroimaging data. As such, it provides a familiar means of reference within MNI reference space. Mapping the model's components onto this reference, then, provides a convenient means to compare the results of a large pool of neuroimaging experiments from a common theoretical framework, a framework that, accounts for a wide range of data from diverse experimental modalities. In other words, it constrains the interpretation of fMRI results with classical lesion data, microstimulation findings, previous functional imaging work, etc. Mapping the model's components to brain anatomy also permits the generation of anatomically explicit simulated

hemodynamic responses based on the model's cell activities. These predictions can then be used to constrain the design and interpretation of functional imaging studies (e.g., Ghosh et al., 2008; Tourville et al., 2008).

Guenther et al. (2006) detailed the neuroanatomical mapping of the model, including a discussion of evidence supporting the assignment of each location. Here we focus on additional assignments given recent modifications of the DIVA model. The MNI locations for recently added regions encompassed by the model are listed in Table 1. These sites are also plotted on a rendering of the SPM canonical brain surface in Figure 2.

Feedforward control

Speech production begins in the model with the activation of a *speech sound map* cell in left premotor and adjacent inferior frontal cortex. According to the model, each frequently encountered speech sound in a speaker's environment is represented by a unique cell in the speech sound map. Cells in the speech sound map project to cells in feedforward *articulator velocity maps* (labeled \dot{M} in Figure 2) in bilateral ventral motor cortex. These projections represent the set of *feedforward motor commands* or articulatory gestures (cf. Browman & Goldstein, 1989) for that speech sound. The feedforward articulator velocity map in each hemisphere consists of eight antagonistic pairs of cells that encode movement velocities for the upper and lower lips, the jaw, the tongue, and the larynx. These velocities ultimately determine the positions of the eight articulators of the Maeda (1990) synthesizer (see *Articulator movements* below for a description of this process). An active speech sound map cell sends a time-varying 16-dimensional input to the feedforward articulator velocity map that encodes the articulator velocities for production of a learned speech sound. The weights are learned during an imitation phase (see below). Feedforward articulator velocity maps are hypothesized to be distributed along the caudal portion of ventrolateral precentral gyrus, the region of primary motor cortex that controls movements of the speech articulators. These cells are hypothesized to correspond to "phasic" cells that have been identified in recordings from motor cortex cells in monkeys (e.g., Kalaska, Cohen, Hyde & Prud'homme, 1989). Ipsilateral and contralateral premotor-to-motor projections that would underlie this hypothesized connectivity have been demonstrated in monkeys (e.g., Dancause, Barbay, Frost, Mahnken & Nudo, 2007; Dancause, Barbay, Frost, Plautz, Popescu et al., 2006; Fang, Stepniewska & Kaas, 2005; Stepniewska, Preuss & Kaas, 2006). Modulation of primary motor cortex by ventral premotor cortex before and during movements has been shown in a number of studies (e.g., Cattaneo, Voss, Brochier, Prabhu, Wolpert et al., 2005; Davare, Lemon & Olivier, 2008). We expect that, in addition to direct premotor to primary motor cortex projections, additional projections via the basal ganglia and/or cerebellum (projecting back to cortex via the thalamus) are also involved in representing the feedforward motor programs.

The mapping from a cell in the speech sound map to the articulator velocity cells is analogous to the process of "phonetic encoding" as conceptualized by Levelt and colleagues (e.g., Levelt, Roelofs & Meyer, 1999; Levelt & Wheeldon, 1994); i.e., it transforms a phonological input from adjacent inferior frontal cortex into the set of feedforward motor commands that produce that sound. The speech sound map, then, can be likened to Levelt et al.'s "mental syllabary," a repository of learned speech motor programs. However, rather than being limited to a repository of motor programs for frequently produced syllables, as Levelt proposed, the speech sound map also represents common syllabic, sub-syllabic (phonemes), and multi-syllabic speech sounds (e.g., words, phrases).

As mentioned above, we hypothesize that this repository of speech motor programs is located in the left hemisphere in right-handed speakers. A dominant role for the left hemisphere has been a hallmark of language processing models for several decades

(Geschwind, 1970). Broca's early findings linking left inferior frontal damage with speech production deficits have been corroborated by a large body of lesion studies (Dronkers, 1996; Duffy, 2005; Hillis, Work, Barker, Jacobs, Breese et al., 2004; Kent & Tjaden, 1997). These clinical findings suggest that speech motor planning is predominantly reliant upon contribution from the posterior inferior frontal region of the left hemisphere. However, the specific level at which the production process becomes lateralized (e.g., semantic, syntactic, phonological, articulatory) has been a topic of debate. We recently showed that the production of single nonsense monosyllables, devoid of semantic or syntactic content, involves left-lateralized contributions from inferior frontal cortex including inferior frontal gyrus, *pars opercularis* (BA 44), ventral premotor, and ventral motor cortex (Ghosh et al., 2008). Activation in these areas is also left-lateralized during monosyllable word production (Tourville et al., 2008). These findings are consistent with the model's assertion that feedforward articulator control originates from cells representing speech motor programs that lie in the left hemisphere.¹ This conclusion, based on neuroimaging data, is consistent with the classical left-hemisphere-dominance view that arose from lesion data. It implies that damage to the left inferior frontal cortex is more commonly associated with speech disruptions than damage to the same region in the right hemisphere because feedforward motor programs are disrupted.

This interpretation is relevant for the study and treatment of speech disorders such as acquired apraxia of speech (AOS). Lesions associated with AOS are predominantly located in the left hemisphere (Duffy, 2005), and particularly affect ventral BA 6 and 44 (ventral precentral gyrus, posterior inferior frontal gyrus, frontal operculum) and the underlying white matter. Our findings corroborate characterizations of AOS as a disruption of the use and development of speech motor programs (Ballard, Granier & Robin, 2000; McNeil, Robin & Schmidt, 2007) and suggest that rehabilitative treatments focused on restoring motor programs, e.g., sound production treatment (Wambaugh, Duffy, McNeil, Robin & Rogers, 2006), and/or improving feedback-based performance should be emphasized.

Feedback control

As indicated in Figure 1, the speech sound map is hypothesized to contribute to both feedforward and feedback control processes. In addition to its projections to the feedforward control map, the speech sound map also projects to *auditory* and *somatosensory target maps*. These projections encode the time-varying sensory expectations, or *targets*, associated with the active speech sound map cell. Auditory targets are given by three pairs of inputs to the auditory target map that describe the upper and lower bounds for the 1st, 2nd, and 3rd formant frequencies of the speech sound being produced. Somatosensory targets consist of a 22-dimensional vector that describes the expected proprioceptive and tactile feedback for the sound being produced. Projections such as these, which predict the sensorimotor state resulting from a movement, are typically described as representing a *forward model* of the movement (e.g. Davidson & Wolpert, 2005; Desmurget & Grafton, 2000; Kawato, 1999; Miall & Wolpert, 1996). According to the model, the auditory and somatosensory target maps send inhibitory inputs to *auditory* and *somatosensory error maps*, respectively. The error maps are effectively the inverse of the target maps: input to the target maps results in inhibition of the region of the error map that represents the expected sensory feedback for the sound being produced. Auditory target and error maps are currently hypothesized to lie in two locations along the posterior superior temporal gyrus. These sites, a lateral one near the superior temporal sulcus, and a medial one at the junction of the temporal and parietal lobes deep in the Sylvian fissure, respond both during speech perception and speech

¹Performance of these motor programs relies on contributions from bilateral premotor, motor, and subcortical regions (see below), areas that are active bilaterally during normal speech production.

production (Buchsbaum, Hickok & Humphries, 2001; Hickok & Poeppel, 2004). In the model, both sites are bilateral. The somatosensory target and state maps lie in ventral supramarginal gyrus, a region Hickok and colleagues (e.g. Hickok & Poeppel, 2004) have argued supports the integration of speech motor commands and sensory feedback. This hypothesized role for ventral parietal cortex during speech production is analogous to the visual-motor integration role associated with more dorsal parietal regions during limb movements (Andersen, 1997; Rizzolatti, Fogassi & Gallese, 1997).

The sensory error maps also receive excitatory inputs from sensory state maps in auditory and somatosensory cortex. The *auditory state map* is hypothesized to lie along Heschl's gyrus and adjacent anterior planum temporale, a region associated with primary and secondary auditory cortex. Cells in the *somatosensory state map* are distributed along the ventral precentral gyrus, roughly mirroring the motor representations on the opposite bank of the central sulcus (see Figure 2). Projections from the auditory and somatosensory state maps relay an estimate of the current sensory state. Activity in the error maps, then, represents the difference between the expected and actual sensory states associated with the production of the current speech sound production.

By effectively “canceling” the self-produced portion of the sensory feedback response, the speech sound map inputs to the sensory target maps function similarly to projections originally described by von Holst and Mittelstaedt (1950) and Sperry (1950). von Holst and Mittelstaedt (1950) proposed the ‘principle of reafference’ in which a copy of the expected sensory consequences of a motor command, termed an *efferece copy*, was subtracted from the realized sensory consequences. A wide body of evidence suggests such a mechanism plays an important role in the motor control of eye and hand movements, as well as speech (e.g. Bays, Flanagan & Wolpert, 2006; Cullen, 2004; Heinks-Maldonado & Houde, 2005; Reppas, Usrey & Reid, 2002; Roy & Cullen, 2004; Voss, Ingram, Haggard & Wolpert, 2006). The hypothesized inhibition of higher order auditory cortex during speech production is supported by several recent studies. Wise and colleagues, using positron emission tomography (PET) to indirectly assess neural activity, noted reduced superior temporal gyrus activation during speech production compared to a listening task (Wise, Greene, Buchel & Scott, 1999). Similarly, comparisons of auditory responses during self-produced speech and while listening to recordings of one's own speech indicate attenuation of auditory cortex responses during speech production (Curio, Neuloh, Numminen, Jousmaki & Hari, 2000; Heinks-Maldonado, Mathalon, Gray & Ford, 2005; Heinks-Maldonado, Nagarajan & Houde, 2006; Numminen, Salmelin & Hari, 1999). Further evidence of auditory response suppression during self-initiated vocalizations is provided by single unit recordings from non-human primates; for example, attenuation of auditory cortical responses prior to self-initiated vocalizations has been demonstrated in the marmoset (Eliades & Wang, 2003, 2005).

If incoming sensory feedback does not fall within the expected target region, an error signal is sent to the *feedback control map* in right frontal/ventral premotor cortex. The feedback control map transforms the auditory and somatosensory error signals into corrective motor velocity commands via projections to the articulator velocity map in bilateral motor cortex. The model's name, DIVA, is an acronym for this mapping from sensory *d*irections *i*nto velocities of *a*rticulators. Feedback-based articulator velocity commands are integrated and combined with the feedforward velocity commands in the articulator position map (see *Articulator movements* below).

The right-lateralized feedback control map was added to the model based on recent results from neuroimaging investigations designed to reveal the neural substrates underlying feedback control of speech production. These studies used fMRI to compare brain activity

during speech production under normal and perturbed auditory (Tourville et al., 2008) and somatosensory (Golfinopoulos, Tourville, Bohland, Ghosh & Guenther, 2009) feedback conditions. Left-lateralized activity in the posterior inferior frontal gyrus *par opercularis*, ventral premotor, and ventral primary motor cortex was noted during the normal feedback condition in both studies. When auditory feedback was perturbed, activity increased bilaterally in the posterior superior temporal cortex, the hypothesized location of the auditory error map, and speakers produced compensatory movements (evident by changes in the acoustic signals produced by the subjects in the perturbed condition compared to the unperturbed condition). The compensatory movements were associated with a *right lateralized* increase in ventral premotor activity. Structural equation modeling (see Tourville et al., 2008 for details) was used to investigate effective connectivity within the network of regions that contributed to feedback-based auditory control. This analysis revealed increased effective connectivity from left posterior temporal cortex to right posterior temporal and ventral premotor cortex (Figure 3). Evidence that right posterior temporal cortex exerts additional influence over motor output via a connection through right inferior frontal gyrus, *par triangularis* (BA 45) during feedback control was also found.

Other imaging studies of speech production that have included a perturbed auditory feedback condition have demonstrated greater right hemisphere involvement in auditory feedback-based control of speech (e.g., Fu, Vythelingum, Brammer, Williams, Amaro et al., 2006; Toyomura, Koyama, Miyamaoto, Terao, Omori et al., 2007). We also noted the same right-lateralized increase of ventral premotor activity associated with auditory feedback perturbation when somatosensory feedback was perturbed (Golfinopoulos et al., 2009). Similarly, a recent study of visuo-motor control reached similar conclusions regarding the relative contributions of the two hemispheres during feedforward- and feedback-based motor control (Grafton, Schmitt, Van Horn & Diedrichsen, 2008). Thus, there is mounting evidence that these hemispheric differences may be a general property of the motor control system.

The implications of lateralized feedforward and feedback motor control of speech may be relevant to the study and treatment of stuttering. Neuroimaging studies of speech production in persons who stutter consistently demonstrate increased right hemisphere activation relative to normal speakers in the precentral and inferior frontal gyrus regions (see Brown, Ingham, Ingham, Laird & Fox, 2005 for review), the same frontal regions were identified as part of the feedback control network by Tourville et al. (2008). It has been hypothesized that stuttering involves excessive reliance upon auditory feedback control due to poor feedforward commands (Max et al., 2004). The current findings provide support for this view: auditory feedback control during the perturbed feedback condition, clearly demonstrated by the behavioral results, was associated with increased activation of right precentral and inferior frontal cortex. According to this view, the right hemisphere inferior frontal activation is a secondary consequence of the root problem, which is aberrant performance in the feedforward system. Poor feedforward performance leads to auditory errors that in turn activate the right-lateralized auditory feedback control system in an attempt to correct for the errors. This hypothesis is consistent with the effects of fluency-inducing therapy on BOLD responses; successful treatment has been associated with a shift toward more normal, left-lateralized frontal activation (De Nil, Kroll, Lafaille & Houle, 2003; Neumann, Preibisch, Euler, von Gudenberg, Lanfermann et al., 2005).

Articulator movement

The feedforward velocity commands and feedback-based error corrective commands are integrated in the *articulator position maps* (labeled *M* in Figure 2) that lie along caudoventral precentral gyrus, adjacent to the feedforward articulator velocity maps. This area is the primary motor representation for muscles of the face and vocal tract. Cells in the articulator

position maps are hypothesized to correspond to “tonic” neurons that have been identified in monkey primary motor cortex (e.g., Kalaska et al., 1989). The map consists of 10 pairs of antagonistic cells² that correspond to parameters of the Maeda vocal tract that determine lip protrusion, upper and lower lip height, jaw height, tongue height, tongue shape, tongue body position, tongue tip location, larynx height, and glottal opening and pressure. Activity in the articulator position maps is a weighted sum of the inputs from the feedforward and feedback-based velocity commands. The relative weight of feedforward and feedback commands in the overall motor command is dependent upon the size of the error signal since this determines the size of the feedback control contribution. The resulting articulator position command drives the simulated vocal tract to produce the desired speech sound.

Based on recent imaging work (Brown, Ngan & Liotti, 2008; Olthoff, Baudewig, Kruse & Dechent, 2008), additional motor cortex cells have been added to the model that represent the intrinsic laryngeal muscles. The locus of a laryngeal representation in motor cortex has typically been associated in the ventrolateral extreme of the precentral gyrus (e.g., Duffy, 2005; Ludlow, 2005), an assumption based heavily upon findings from non-human primates (e.g. Simonyan & Jurgens, 2003) and supported by the intracortical mapping studies performed by Penfield and colleagues in humans prior to epilepsy surgery (Penfield & Rasmussen, 1950; Penfield & Roberts, 1959). Accordingly, a motor larynx representation in this location was included in our initial anatomical mapping of the model (Guenther et al., 2006). Brown et al. (2008) and Orloff et al. (2008) have since demonstrated a bilateral representation in a more dorsal region of motor cortex adjacent to the lip area and near a second “vocalization” region identified by Penfield and Roberts (1959, p. 200). The authors also noted a ventral representation near/within bilateral Rolandic operculum that is consistent with the non-human primate literature. Brown and colleagues (2008) concluded that the dorsal region likely represents the intrinsic laryngeal muscles that control the size of the glottal opening. The opercular representation, it was speculated, likely represents the extrinsic laryngeal muscles that affect vocal tract resonances by controlling larynx height. Based on these findings, two sets of cells representing laryngeal parameters of the Maeda articulator model (Maeda, 1990) associated with laryngeal functions have been assigned to MNI locations: cells in ventrolateral precentral gyrus (labeled *Larynx, Extrinsic* in Table 1) represent larynx height, whereas cells in the dorsomedial orofacial region of precentral gyrus (labeled *Larynx, Intrinsic* in Figure 2) represent a weighted sum of parameters representing glottal opening and glottal pressure.

Speech production is consistently associated with bilateral activity in the medial prefrontal cortex, including the supplementary motor area (SMA), in the basal ganglia, and in the thalamus (e.g., Bohland & Guenther, 2006; Ghosh et al., 2008; Tourville et al., 2008). Previous versions of the DIVA model have offered no account of this activity. The SMA is strongly interconnected with lateral motor and premotor cortex and the basal ganglia (Jurgens, 1984; Lehericy, Ducros, Krainik, Francois, Van de Moortele et al., 2004; Luppino, Matelli, Camarda & Rizzolatti, 1993; Matsumoto, Nair, LaPresto, Bingaman, Shibasaki et al., 2007; Matsumoto, Nair, LaPresto, Najm, Bingaman et al., 2004). Recordings from primates have revealed cells in the SMA that encode movement dynamics (Padoa-Schioppa, Li & Bizzi, 2004). Cells representing higher-order information regarding the planning/performance of sequences of movements have also been identified in the SMA. Activity representing particular sequences of movements to be performed (Shima & Tanji, 2000), intervals between specific movements within a sequence (Shima & Tanji, 2000), the ordinal position of movements within a sequence (Clower & Alexander, 1998), and the number of items remaining in a sequence (Sohn & Lee, 2007) have been noted in recordings

²The activity of cells representing glottal opening and pressure are not determined dynamically via input from the feedforward and feedback velocity commands. Their activity is given by a weighted sum of Maeda vocal tract parameters that are set manually.

of neurons in the SMA. Microstimulation of the SMA in humans yields vocalization, word or syllable repetitions, and/or speech arrest (Penfield & Welch, 1951). Bilateral damage to these areas results in speech production deficits including transcortical motor aphasia (Jonas, 1981; Ziegler, Kilian & Deger, 1997) and akinetic mutism (Adams, 1989; Mochizuki & Saito, 1990; Nemeth, Hegedus & Molnar, 1988).

These data and recent imaging findings have led a number of investigators to conclude that the SMA plays a critical role in controlling the initiation of speech motor commands (e.g., Alario, Chainay, Lehericy & Cohen, 2006; Bohland & Guenther, 2006; Jonas, 1987; Ziegler et al., 1997). The SMA is reciprocally connected with the basal ganglia, another region widely believed to contribute to gating motor commands (e.g., Albin, Young & Penney, 1995; Pickett, Kuniholm, Protopapas, Friedman & Lieberman, 1998; Van Buren, 1963). The basal ganglia receive afferents from most areas of the cerebral cortex, including motor and prefrontal regions and, notably, associative and limbic cortices. Thus, the basal ganglia are well-suited for integrating contextual cues for the purpose of gating motor commands.

Based on these findings, an *initiation map*, hypothesized to lie in the SMA, has been added to the DIVA model. The initiation map gates the release articulator position commands to the periphery. According to the model, each speech motor program in the speech sound map is associated with a cell in the initiation map. Motor commands associated with that program are released when the corresponding initiation map cell becomes active. Activity in the initiation map (I) is given by:

$$I_i(t) = 1 \text{ if the } i^{\text{th}} \text{ sound is being produced or perceived}$$

$$I_i(t) = 0 \text{ otherwise}$$

The timing of initiation cell activity is governed by contextual inputs from the basal ganglia via the thalamus. Presently, this timing is simply based on a delayed input from the speech sound map. Cells representing the initiation map have been placed bilaterally in the SMA, caudate, putamen, globus pallidus and thalamus, all of which demonstrate activity during simple speech production tasks (e.g., Ghosh et al., 2008; Tourville et al., 2008). A model of speech motor sequence planning and execution, the GODIVA model, has been developed (Bohland, Bullock & Guenther, in press; Bohland & Guenther, 2006) that provides a comprehensive account of the interactions between the SMA, basal ganglia, motor, and premotor cortex that result in the gating of speech motor commands; due to space limitations we refer the interested reader to that publication for details.

LEARNING IN THE DIVA MODEL

Early babbling phase

Before DIVA is able to produce speech sounds, the mappings between the various components of the model must be learned. The model first learns the relationship between motor commands and their sensory consequences during a process analogous to infant babbling. The mappings that are tuned in this process are highlighted in the simplified DIVA block diagram shown in Figure 4. For clarity, anatomical labels have been removed and the names of the model's components have been shortened. During the babbling phase, pseudo-random articulator movements provide auditory and somatosensory feedback that is compared to the causative motor commands. The paired motor and sensory information is used to tune the synaptic projections from temporal and parietal sensory error maps to the feedforward control map. Once tuned, these projections transform sensory error signals into corrective motor velocity commands. This mapping from desired sensory outcome to the appropriate motor action is an inverse kinematic transformation and is often referred to as an *inverse model* (e.g., Kawato, 1999; Wolpert & Kawato, 1998).

The cerebellum is a likely contributor to the feedback motor command. Neuroimaging studies of motor learning have noted cerebellar activity that is associated with the size or frequency of sensory error (e.g., Blakemore, Frith & Wolpert, 2001; Blakemore, Wolpert & Frith, 1999; Diedrichsen, Hashambhoy, Rane & Shadmehr, 2005; Flament, Ellermann, Kim, Ugurbil & Ebner, 1996; Grafton et al., 2008; Imamizu, Higuchi, Toda & Kawato, 2007; Imamizu, Miyauchi, Tamada, Sasaki, Takino et al., 2000; Miall & Jenkinson, 2005; Schreurs, McIntosh, Bahro, Herscovitch, Sunderland et al., 1997; Tesche & Karhu, 2000). It has been speculated that a representation of sensory errors in the cerebellum drives corrective motor commands (Grafton et al., 2008; Penhune & Doyon, 2005) and contributes to feedback-based motor learning (Ito, 2000; Tseng, Diedrichsen, Krakauer, Shadmehr & Bastian, 2007; Wolpert, Miall & Kawato, 1998). For instance, the cerebellum has been hypothesized to support the learning of inverse kinematics (e.g., Kawato, 1999; Wolpert & Kawato, 1998), a role for which it is anatomically well-suited: the cerebellum receives inputs from higher-order auditory and somatosensory areas (e.g., Schmahmann & Pandya, 1997), and projects heavily to the motor cortex (Middleton & Strick, 1997). Based on the cerebellum's putative role in feedback-based motor learning, it is hypothesized to contribute to the mapping between sensory states and motor cortex, i.e., the projections that encode the feedback motor command.

Imitation phase

Feedback control system—Once the general sensory-to-motor mapping described above has been learned, the model undergoes a second learning phase that is specific to the production of speech sounds. This phase can be subdivided into two components. In the first component, weights from the speech sound map are tuned. Analogous to the exposure an infant has to the sounds of his/her native language, the model is presented speech sound samples (e.g., phonemes, syllables, words). The speech samples take the form of time varying acoustic signals spoken by a human speaker. According to the model, when a new speech sound is presented, it becomes associated with an unused cell in the inferior frontal speech sound map (via temporo-frontal projections not shown in Figure 2). With subsequent exposures to that speech sound, the model learns an auditory target for that sound in the form of a time-varying region that encodes the allowable variability in the acoustic signal (see Guenther, 1995, for a description of the learning laws that govern this process). During the second component of learning in the feedback control system, weights from the speech sound map to the somatosensory target map are tuned during correct self productions.

Reciprocal pathways between inferior frontal and auditory and somatosensory cortex have been demonstrated in humans (Makris, Kennedy, McInerney, Sorensen, Wang et al., 2005; Matsumoto et al., 2004) and non-human primates (Morel & Kaas, 1992; Ojemann, 1991; Romanski, Tian, Fritz, Mishkin, Goldman-Rakic et al., 1999; Schmahmann & Pandya, 2006); also see Duffau (2008) for a description of the putative fronto-parietal and fronto-temporal pathways involved in language processing. It has been argued by many that the cerebellum uses sensory error to build forward models that generate sensory predictions (Blakemore et al., 2001; Imamizu et al., 2000; Kawato, Kuroda, Imamizu, Nakano, Miyauchi et al., 2003; O'Reilly, Mesulam & Nobre, 2008), the role of projections from the speech sound map to sensory target maps in DIVA model. It is likely, therefore, that the cerebellum contributes to the attenuation of sensory target representation in sensory cortex (cf. Blakemore et al., 2001). For this reason, cerebellar side loops are hypothesized in the projections from the speech sound map to the sensory target maps.

Feedforward control system—Feedforward commands are also learned during the imitation phase, once auditory targets have been learned. Initial attempts to produce the speech sound result in large sensory error signals due to poorly tuned projections from the

speech sound map cells to the primary motor cortex articulatory velocity and position maps, and production relies heavily on the feedback control system. With each production, however, the feedback-based corrective motor command is added to the weights from the speech sound map to the feedforward articulator velocity cells, incrementally improving the accuracy of the feedforward motor command. With practice, the feedforward commands become capable of driving production of the speech sound with minimal sensory error and, therefore, little reliance on the feedback control system unless unexpected sensory feedback is encountered (e.g., due to changing vocal tract dynamics, a bite block, or artificial auditory feedback perturbation).

It is widely held that the cerebellum is involved with the learning and maintenance of feedforward motor commands (though see Grafton et al., 2008; Kawato, 1999; Ohyama, Nores, Murphy & Mauk, 2003). The cerebellum receives input from premotor, auditory, and somatosensory cortical areas via the pontine nuclei and projects heavily to back to motor cortex via the ventral thalamus (Middleton & Strick, 1997). This circuitry provides a substrate for the integration of sensory state information that may be important for choosing motor commands (e.g., Schmahmann & Pandya, 1997), and projects heavily to the motor cortex. In the DIVA model, the cerebellum is therefore included as a side loop in the projection from the speech sound map to the articulator velocity map.

Lesions to anterior vermal and paravermal cerebellum have been associated with disruptions of speech production (Ackermann, Vogel, Petersen & Poremba, 1992; Urban, Marx, Hunsche, Gawehn, Vucurevic et al., 2003), termed ataxic dysarthria, characterized by an impaired ability to produce speech with fluent timing and gestural coordination. This region is typically active bilaterally during overt speech production in neuroimaging experiments. Additional activity is typically found in adjacent lateral cortex bilaterally (Bohland & Guenther, 2006; Ghosh et al., 2008; Riecker, Ackermann, Wildgruber, Dogil & Grodd, 2000; Riecker, Wildgruber, Dogil, Grodd & Ackermann, 2002; Tourville et al., 2008; Wildgruber, Ackermann & Grodd, 2001), an area less commonly associated with ataxic dysarthria. Model cells have therefore been placed bilaterally in two cerebellar cortical regions: anterior paravermal cortex (not visible in Figure 2) and superior lateral cortex (*Lat. Cbm*). The former are part of the feedforward control system, while the latter are hypothesized to contribute to the sensory predictions that form the auditory and somatosensory targets in the feedback control system.

The speech sound map and mirror neurons

The role played by the speech sound map in the DIVA model is similar to that attributed to “mirror neurons” (Kohler, Keysers, Umiltà, Fogassi, Gallese et al., 2002; Rizzolatti, Fadiga, Gallese & Fogassi, 1996), so termed because they respond both while performing an action and perceiving an action. Mirror neurons in non-human primates have been shown to code for complex actions such as grasping rather than the individual movements that comprise an action (Rizzolatti, Camarda, Fogassi, Gentilucci, Luppino et al., 1988). Neurons within the speech sound map are hypothesized to embody similar properties: activation during speech production drives complex articulator movement via projections to articulator velocity cells in motor cortex, and activation during speech perception tunes connections between the speech sound map and sensory target maps in auditory and somatosensory cortex. Evidence of mirror neurons in humans has implicated left precentral gyrus for grasping actions (Tai, Scherfler, Brooks, Sawamoto & Castiello, 2004), and left opercular inferior frontal gyrus for finger movements (Iacoboni, Woods, Brass, Bekkering, Mazziotta et al., 1999). Mirror neurons related to communicative mouth movements have been found in monkey area F5 (Ferrari, Gallese, Rizzolatti & Fogassi, 2003) immediately lateral to their location for grasping movements (di Pellegrino, Fadiga, Fogassi, Gallese & Rizzolatti, 1992). This area has been proposed to correspond to the caudal portion of ventral inferior frontal gyrus

(Brodmann's area 44) in the human (see Binkofski & Buccino, 2004; Rizzolatti & Arbib, 1998).

CURRENT PERSPECTIVES

The DIVA model provides a computationally explicit account of the interactions between the brain regions involved in speech acquisition and production. The model has proven to be a valuable tool for studying the mechanisms underlying normal (Callan et al., 2000; Lane et al., 2007; Perkell et al., 2007; Perkell et al., 2000; Perkell et al., 2004; Perkell et al., 2004; Villacorta, Perkell & Guenther, 2007) and disordered speech (Max et al., 2004; Robin et al., 2008; Terband et al., 2008). Because the model is expressed as a neural network, it provides a convenient substrate for generating predictions that are well-suited for empirical testing. Importantly, the model's development has been constrained to biologically plausible mechanisms. Thus, as DIVA has come to account for a wide range of speech production phenomena (e.g., Callan et al., 2000; Guenther, 1994; Guenther, 1995; Guenther et al., 1998; Nieto-Castanon et al., 2005), it does so from a unified *quantitative* and *neurobiologically grounded* framework.

In this article, we have reviewed the key elements of the DIVA model, focusing on recent developments based on results from functional imaging experiments. Feedforward and feedback control maps are hypothesized to lie in left and right ventral premotor cortex, respectively. The lateralized motor control mechanisms embodied by the model may provide useful insight into the study and treatment of speech disorders. Questions remain, however, regarding the interaction between lateralized frontal and largely bilateral sensory processes. As discussed above, the data are consistent with DIVA-predicted projections from left-lateralized premotor cells to bilateral auditory cortex that encode sensory expectations. A further prediction of the model is that clear and stressed speech involve the use of smaller, more precise sensory targets compared to normal or fast speech (Guenther, 1995). If correct, increased activity should be seen in auditory and somatosensory cortical areas during clear and stressed speaking conditions, corresponding to increased error cell activity due to the more precise sensory targets. This prediction is currently being tested in an ongoing fMRI experiment.

Our experimental data also suggest that projections from bilateral auditory cells to right premotor cortex are involved in transforming auditory errors into corrective motor commands. The anatomical pathways that support these mechanisms are not fully understood. Further study of the information conveyed by those projections is also necessary. Studies have begun to explore putative sensory expectation projections from lateral frontal cortex to auditory cortex, establishing an inhibitory effect linked to ongoing articulator movements (e.g., Heinks-Maldonado et al., 2006). A clear understanding of the units of this inhibitory input (e.g., does it have an acoustic, articulatory, or phonological organization), as well as the error maps themselves, is yet to be fully established. Functional imaging experiments focused on these questions are currently underway.

The model has also expanded to include representations of the supplementary motor area and basal ganglia, which are hypothesized to provide a gating signal that initiates the release of motor commands to the speech articulators. In its current form, this *initiation map* is highly simplified. Mechanisms for learning the appropriate timing of motor command release have not yet been incorporated into the model. The brain regions associated with the model's initiation map, along with the pre-supplementary motor area (e.g., Clower & Alexander, 1998; Shima & Tanji, 2000), have also been implicated in the selection and proper sequencing of individual motor programs for serial production. Bohland and Guenther (2006) investigated the brain regions that contribute to the assembly and

performance of speech sound sequences. A neural network model of the mechanisms underlying this process, including interactions between the various cortical and subcortical brain regions involved, has been developed (Bohland et al., in press). Outputs from this speech sequencing model, termed GODIVA, serve as inputs to the DIVA model's speech sound map. This work bridges the gap between computational models of the linguistic/phonological level of speech production (Dell, 1986; Hartley & Houghton, 1996; Levelt et al., 1999), and DIVA, which addresses production at the speech motor control level. Like the DIVA model, GODIVA is neurobiologically plausible and thus represents an expanded substrate for the study of normal and disorder speech processing.

An important aspect of speech production not addressed in the DIVA model is the control of prosody. Modulation of pitch, loudness, duration, and rhythm convey meaningful linguistic and affective cues (Bolinger, 1961, 1989; Lehiste, 1970, 1976; Netsell, 1973; Shriberg & Kent, 1982). The DIVA model has addressed speech motor control as the segmental level (phoneme or syllable units). With the development of the GODIVA model, we have started to account for speech production at a suprasegmental level. We have recently begun a similar expansion of the model to allow for control of prosodic cues, which often operates over multiple individual segments. Experiments currently under way are investigating whether the various prosodic features (volume, duration, and pitch) are controlled independently to achieve a desired stress level or if a combined "stress target" is set that is reached by a dynamic combination of individual features. We are testing these alternative hypotheses measuring speaker compensations to perturbations of pitch and loudness. Adaptive responses limited to the perturbed modality supports the notion that prosodic features are independently control; adaptation across modalities is evidence for an integrated "stress" controller.

As we have done in the past, we intend to couple our modeling efforts with investigations of the neural bases of prosodic control. There is agreement in the literature that no single brain region is responsible for prosodic control, but there is little consensus regarding which regions are involved and in what capacity (see Sidtis & Van Lancker Sidtis, 2003, for review). One of the more consistent findings in the literature concerns perception and production of affective prosody, which appear to rely more on the right cerebral hemisphere than the left hemisphere (Adolphs, Damasio & Tranel, 2002; Buchanan, Lutz, Mirzazade, Specht, Shah et al., 2000; George, Parekh, Rosinsky, Ketter, Kimbrell et al., 1996; Ghacibeh & Heilman, 2003; Kotz, Meyer, Alter, Besson, von Cramon et al., 2003; Mitchell, Elliott, Barry, Cruttenden & Woodruff, 2003; Pihan, Altenmuller & Ackermann, 1997; Ross & Mesulam, 1979; Williamson, Harrison, Shenal, Rhodes & Demaree, 2003), though the view of affective prosody as a unitary, purely right hemisphere entity is oversimplified (Sidtis & Van Lancker Sidtis, 2003) and there is considerable debate over which and to what extent prosodic feature control is lateralized (e.g., Doherty, West, Dilley, Shattuck-Hufnagel & Caplan, 2004; Emmorey, 1987; Gandour, Dziedzic, Wong, Lowe, Tong et al., 2003; Meyer, Alter, Friederici, Lohmann & von Cramon, 2002; Stiller, Gaschler-Markefski, Baumgart, Schindler, Tempelmann et al., 1997; Walker, Pelletier & Reif, 2004). We hope to clarify this understanding by comparing the neural responses associated with prosodic control to those involved in formant control as indicated by our formant perturbation imaging study (Tourville et al., 2008). Our focus is on differences in the two control networks, particularly the laterality of sensory and motor cortical response. With this effort we hope to continue our progress toward building a comprehensive, unified account of the neural mechanisms underlying speech motor control.

Acknowledgments

Supported by the National Institute on Deafness and other Communication Disorders (R01 DC02852, F. Guenther PI).

REFERENCES

- Ackermann H, Vogel M, Petersen D, Poremba M. Speech deficits in ischaemic cerebellar lesions. *Journal of Neurology*. 1992; 239(4):223–227. [PubMed: 1597689]
- Adams, RD. *Principles of Neurology*. New York: McGraw-Hill; 1989.
- Adolphs R, Damasio H, Tranel D. Neural systems for recognition of emotional prosody: a 3-D lesion study. *Emotion*. 2002; 2(1):23–51. [PubMed: 12899365]
- Alario FX, Chainay H, Lehericy S, Cohen L. The role of the supplementary motor area (SMA) in word production. *Brain Research*. 2006; 1076(1):129–143. [PubMed: 16480694]
- Albin RL, Young AB, Penney JB. The functional anatomy of disorders of the basal ganglia. *Trends in Neurosciences*. 1995; 18(2):63–64. [PubMed: 7537410]
- Andersen RA. Multimodal integration for the representation of space in the posterior parietal cortex. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*. 1997; 352(1360):1421–1428.
- Ballard KJ, Granier JP, Robin DA. Understanding the nature of apraxia of speech: Theory, analysis, and treatment. *Aphasiology*. 2000; 14(10):969–995.
- Bays PM, Flanagan JR, Wolpert DM. Attenuation of self-generated tactile sensations is predictive, not postdictive. *Public Library of Science Biology*. 2006; 4(2):e28.
- Belliveau JW, Kwong KK, Kennedy DN, Baker JR, Stern CE, Benson R, et al. Magnetic resonance imaging mapping of brain function. Human visual cortex. *Investigative Radiology*. 1992; 27(Suppl 2):S59–S65. [PubMed: 1468876]
- Binkofski F, Buccino G. Motor functions of the Broca's region. *Brain and Language*. 2004; 89(2):362–369. [PubMed: 15068919]
- Blakemore SJ, Frith CD, Wolpert DM. The cerebellum is involved in predicting the sensory consequences of action. *Neuroreport*. 2001; 12(9):1879–1884. [PubMed: 11435916]
- Blakemore SJ, Wolpert DM, Frith CD. The cerebellum contributes to somatosensory cortical activity during self-produced tactile stimulation. *Neuroimage*. 1999; 10(4):448–459. [PubMed: 10493902]
- Bohland JW, Bullock D, Guenther FH. Neural representations and mechanisms for the performance of simple speech sequences. *Journal of Cognitive Neuroscience*. (in press).
- Bohland JW, Guenther FH. An fMRI investigation of syllable sequence production. *Neuroimage*. 2006; 32:821–841. [PubMed: 16730195]
- Browman CP, Goldstein L. Articulatory gestures as phonological units. *Phonology*. 1989; 6:201–251.
- Brown S, Ingham RJ, Ingham JC, Laird AR, Fox PT. Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies. *Human Brain Mapping*. 2005; 25(1):105–117. [PubMed: 15846815]
- Brown S, Ngan E, Liotti M. A larynx area in the human motor cortex. *Cerebral Cortex*. 2008; 18(4): 837–845. [PubMed: 17652461]
- Buchanan TW, Lutz K, Mirzazade S, Specht K, Shah NJ, Zilles K, et al. Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Brain Res Cogn Brain Res*. 2000; 9(3):227–238. [PubMed: 10808134]
- Buchsbaum BR, Hickok G, Humphries C. Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. *Cognitive Science*. 2001; 25(5): 663–678.
- Callan DE, Kent RD, Guenther FH, Vorperian HK. An auditory-feedback-based neural network model of speech production that is robust to developmental changes in the size and shape of the articulatory system. *Journal of Speech, Language, and Hearing Research*. 2000; 43(3):721–736.
- Cattaneo L, Voss M, Brochier T, Prabhu G, Wolpert DM, Lemon RN. A cortico-cortical mechanism mediating object-driven grasp in humans. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102(3):898–903. [PubMed: 15642941]

- Clower WT, Alexander GE. Movement sequence-related activity reflecting numerical order of components in supplementary and presupplementary motor areas. *Journal of Neurophysiology*. 1998; 80(3):1562–1566. [PubMed: 9744961]
- Cullen KE. Sensory signals during active versus passive movement. *Current Opinions in Neurobiology*. 2004; 14(6):698–706.
- Curio G, Neuloh G, Numminen J, Jousmaki V, Hari R. Speaking modifies voice-evoked activity in the human auditory cortex. *Human Brain Mapping*. 2000; 9(4):183–191. [PubMed: 10770228]
- Dancause N, Barbay S, Frost SB, Mahnken JD, Nudo RJ. Interhemispheric connections of the ventral premotor cortex in a new world primate. *Journal of Comparative Neurology*. 2007; 505(6):701–715. [PubMed: 17948893]
- Dancause N, Barbay S, Frost SB, Plautz EJ, Popescu M, Dixon PM, et al. Topographically divergent and convergent connectivity between premotor and primary motor cortex. *Cerebral Cortex*. 2006; 16(8):1057–1068. [PubMed: 16221929]
- Davare M, Lemon R, Olivier E. Selective modulation of interactions between ventral premotor cortex and primary motor cortex during precision grasping in humans. *Journal of Physiology-London*. 2008; 586(11):2735–2742.
- Davidson PR, Wolpert DM. Widespread access to predictive models in the motor system: a short review. *Journal of Neural Engineering*. 2005; 2(3):S313–S319. [PubMed: 16135891]
- De Nil LF, Kroll RM, Lafaille SJ, Houle S. A positron emission tomography study of short- and long-term treatment effects on functional brain activation in adults who stutter. *Journal of Fluency Disorders*. 2003; 28(4):357–379. quiz 379-80. [PubMed: 14643070]
- Dell GS. A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review*. 1986; 93(3):283–321. [PubMed: 3749399]
- Desmurget M, Grafton S. Forward modeling allows feedback control for fast reaching movements. *Trends in Cognitive Science*. 2000; 4(11):423–431.
- di Pellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G. Understanding motor events: a neurophysiological study. *Experimental Brain Research*. 1992; 91(1):176–180.
- Diedrichsen J, Hashambhoy Y, Rane T, Shadmehr R. Neural correlates of reach errors. *Journal of Neuroscience*. 2005; 25(43):9919–9931. [PubMed: 16251440]
- Doherty CP, West WC, Dilley LC, Shattuck-Hufnagel S, Caplan D. Question/statement judgments: an fMRI study of intonation processing. *Human Brain Mapping*. 2004; 23(2):85–98. [PubMed: 15340931]
- Dronkers NF. A new brain region for coordinating speech articulation. *Nature*. 1996; 384(6605):159–161. [PubMed: 8906789]
- Duffau H. The anatomo-functional connectivity of language revisited new insights provided by electrostimulation and tractography. *Neuropsychologia*. 2008; 46(4):927–934. [PubMed: 18093622]
- Duffy, JR. *Motor speech disorders: Substrates, differential diagnosis, and management*. St. Louis, MO: Elsevier Mosby; 2005.
- Eliades SJ, Wang X. Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *Journal of Neurophysiology*. 2003; 89(4):2194–2207. [PubMed: 12612021]
- Eliades SJ, Wang X. Dynamics of auditory-vocal interaction in monkey auditory cortex. *Cerebral Cortex*. 2005; 15(10):1510–1523. [PubMed: 15689521]
- Emmorey KD. The neurological substrates for prosodic aspects of speech. *Brain and Language*. 1987; 30(2):305–320. [PubMed: 3567552]
- Fang PC, Stepniewska I, Kaas JH. Ipsilateral cortical connections of motor, premotor, frontal eye, posterior parietal fields in a prosimian primate, *Otolemur garnetti*. *Journal of Comparative Neurology*. 2005; 490(3):305–333. [PubMed: 16082679]
- Ferrari PF, Gallese V, Rizzolatti G, Fogassi L. Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience*. 2003; 17(8):1703–1714. [PubMed: 12752388]
- Fiez JA, Petersen SE. Neuroimaging studies of word reading. *Proceedings of the National Academy of Sciences of the United States of America*. 1998; 95(3):914–921. [PubMed: 9448259]

- Flament D, Ellermann JM, Kim SG, Ugurbil K, Ebner TJ. Functional magnetic resonance imaging of cerebellar activation during the learning of visuomotor dissociation task. *Human Brain Mapping*. 1996; 4(3):210–226. [PubMed: 20408199]
- Friston KJ, Holmes AP, Poline JB, Grasby PJ, Williams SC, Frackowiak RS, et al. Analysis of fMRI time-series revisited. *Neuroimage*. 1995; 2(1):45–53. [PubMed: 9343589]
- Fu CH, Vythelingum GN, Brammer MJ, Williams SC, Amaro E Jr, Andrew CM, et al. An fMRI study of verbal self-monitoring: neural correlates of auditory verbal feedback. *Cerebral Cortex*. 2006; 16(7):969–977. [PubMed: 16195470]
- Gandour J, Dziedzic M, Wong D, Lowe M, Tong Y, Hsieh L, et al. Temporal integration of speech prosody is shaped by language experience: an fMRI study. *Brain and Language*. 2003; 84(3):318–336. [PubMed: 12662974]
- George MS, Parekh PI, Rosinsky N, Ketter TA, Kimbrell TA, Heilman KM, et al. Understanding emotional prosody activates right hemisphere regions. *Archives of Neurology*. 1996; 53(7):665–670. [PubMed: 8929174]
- Geschwind N. The organization of language and the brain. *Science*. 1970; 170(961):940–944. [PubMed: 5475022]
- Ghacibeh GA, Heilman KM. Progressive affective aprosodia and prosoplegia. *Neurology*. 2003; 60(7):1192–1194. [PubMed: 12682334]
- Ghosh SS, Bohland JW, Guenther FH. Comparisons of brain regions involved in overt production of elementary phonetic units. *Neuroimage (9th Meeting of the Organization for Human Brain Mapping, New York)*. 2003; 19(2):S57.
- Ghosh SS, Tourville JA, Guenther FH. An fMRI study of the overt production of simple speech sounds. *Journal of Speech, Language, and Hearing Research*. 2008; 51:1183–1202.
- Goense JB, Logothetis NK. Neurophysiology of the BOLD fMRI signal in awake monkeys. *Current Biology*. 2008; 18(9):631–640. [PubMed: 18439825]
- Golfinopoulos E, Tourville JA, Bohland JW, Ghosh SS, Guenther FH. Neural circuitry underlying somatosensory feedback control of speech production. 2009 In preparation.
- Grafton ST, Schmitt P, Van Horn J, Diedrichsen J. Neural substrates of visuomotor learning based on improved feedback and prediction. *Neuroimage*. 2008; 39:1383–1395. [PubMed: 18032069]
- Guenther FH. A neural network model of speech acquisition and motor equivalent speech production. *Biological Cybernetics*. 1994; 72(1):43–53. [PubMed: 7880914]
- Guenther FH. Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*. 1995; 102(3):594–621. [PubMed: 7624456]
- Guenther FH, Espy-Wilson CY, Boyce SE, Matthies ML, Zandipour M, Perkell JS. Articulatory tradeoffs reduce acoustic variability during American English /r/ production. *Journal of the Acoustical Society of America*. 1999; 105(5):2854–2865. [PubMed: 10335635]
- Guenther FH, Ghosh SS, Tourville JA. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*. 2006; 96(3):280–301. [PubMed: 16040108]
- Guenther FH, Hampson M, Johnson D. A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review*. 1998; 105(4):611–633. [PubMed: 9830375]
- Hartley T, Houghton G. A linguistically constrained model of short-term memory for nonwords. *Journal of Memory and Language*. 1996; 35(1):1–31.
- Heinks-Maldonado TH, Houde JF. Compensatory responses to brief perturbations of speech amplitude. *Acoustics Research Letters Online*. 2005; 6(3):131–137.
- Heinks-Maldonado TH, Mathalon DH, Gray M, Ford JM. Fine-tuning of auditory cortex during speech production. *Psychophysiology*. 2005; 42(2):180–190. [PubMed: 15787855]
- Heinks-Maldonado TH, Nagarajan SS, Houde JF. Magnetoencephalographic evidence for a precise forward model in speech production. *Neuroreport*. 2006; 17(13):1375–1379. [PubMed: 16932142]
- Hickok G, Poeppel D. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*. 2004; 92(1–2):67–99. [PubMed: 15037127]
- Hillis AE, Work M, Barker PB, Jacobs MA, Breese EL, Maurer K. Re-examining the brain regions crucial for orchestrating speech articulation. *Brain*. 2004; 127(Pt 7):1479–1487. [PubMed: 15090478]

- Jacoboni M, Woods RP, Brass M, Bekkering H, Mazziotta JC, Rizzolatti G. Cortical mechanisms of human imitation. *Science*. 1999; 286(5449):2526–2528. [PubMed: 10617472]
- Imamizu H, Higuchi S, Toda A, Kawato M. Reorganization of brain activity for multiple internal models after short but intensive training. *Cortex*. 2007; 43(3):338–349. [PubMed: 17533757]
- Imamizu H, Miyauchi S, Tamada T, Sasaki Y, Takino R, Putz B, et al. Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature*. 2000; 403(6766):192–195. [PubMed: 10646603]
- Indefrey P, Levelt WJ. The spatial and temporal signatures of word production components. *Cognition*. 2004; 92(1–2):101–144. [PubMed: 15037128]
- Ito M. Mechanisms of motor learning in the cerebellum. *Brain Research*. 2000; 886(1–2):237–245. [PubMed: 11119699]
- Jonas S. The supplementary motor region and speech emission. *Journal of Communication Disorders*. 1981; 14:349–373. [PubMed: 7287912]
- Jonas, S. The supplementary motor region and speech. In: Perecman, E., editor. *The frontal lobes revisited*. New York: IRBN Press; 1987. p. 241-250.
- Jurgens U. The efferent and efferent connections of the supplementary motor area. *Brain Research*. 1984; 300:63–81. [PubMed: 6733468]
- Kalaska JF, Cohen DA, Hyde ML, Prud'homme M. A comparison of movement direction-related versus load direction-related activity in primate motor cortex, using a two-dimensional reaching task. *Journal of Neuroscience*. 1989; 9(6):2080–2102. [PubMed: 2723767]
- Kawato M. Internal models for motor control and trajectory planning. *Current Opinions in Neurobiology*. 1999; 9(6):718–727.
- Kawato M, Kuroda T, Imamizu H, Nakano E, Miyauchi S, Yoshioka T. Internal forward models in the cerebellum: fMRI study on grip force and load force coupling. *Progress in Brain Research*. 2003; 142:171–188. [PubMed: 12693261]
- Kent, RD.; Tjaden, K. Brain functions underlying speech. In: Hardcastle, WJ.; Laver, J., editors. *Handbook of phonetic sciences*. Oxford: Blackwell; 1997. p. 220-255.
- Kohler E, Keysers C, Umiltà MA, Fogassi L, Gallese V, Rizzolatti G. Hearing sounds, understanding actions: action representation in mirror neurons. *Science*. 2002; 297(5582):846–848. [PubMed: 12161656]
- Kotz SA, Meyer M, Alter K, Besson M, von Cramon DY, Friederici AD. On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain and Language*. 2003; 86(3):366–376. [PubMed: 12972367]
- Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Weisskoff RM, Poncelet BP, et al. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences of the United States of America*. 1992; 89(12):5675–5679. [PubMed: 1608978]
- Lane H, Denny M, Guenther FH, Hanson HM, Marrone N, Matthies ML, et al. On the structure of phoneme categories in listeners with cochlear implants. *Journal of Speech, Language, and Hearing Research*. 2007; 50(1):2–14.
- Lane H, Denny M, Guenther FH, Matthies ML, Menard L, Perkell JS, et al. Effects of bite blocks and hearing status on vowel production. *J Acoust Soc Am*. 2005; 118(3 Pt 1):1636–1646. [PubMed: 16240823]
- Lane H, Matthies ML, Guenther FH, Denny M, Perkell JS, Stockmann E, et al. Effects of short- and long-term changes in auditory feedback on vowel and sibilant contrasts. *Journal of Speech, Language, and Hearing Research*. 2007; 50(4):913–927.
- Lehericy S, Ducros M, Krainik A, Francois C, Van de Moortele PF, Ugurbil K, et al. 3-D diffusion tensor axonal tracking shows distinct SMA and pre-SMA projections to the human striatum. *Cerebral Cortex*. 2004; 14(12):1302–1309. [PubMed: 15166103]
- Levelt WJ, Roelofs A, Meyer AS. A theory of lexical access in speech production. *Behavioral and Brain Sciences*. 1999; 22(1):1–38. discussion 38–75. [PubMed: 11301520]
- Levelt WJ, Wheeldon L. Do speakers have access to a mental syllabary? *Cognition*. 1994; 50(1–3):239–269. [PubMed: 8039363]

- Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. *Nature*. 2001; 412(6843):150–157. [PubMed: 11449264]
- Ludlow CL. Central nervous system control of the laryngeal muscles in humans. *Respiratory Physiology and Neurobiology*. 2005; 147(2–3):205–222. [PubMed: 15927543]
- Luppino G, Matelli M, Camarda R, Rizzolatti G. Corticocortical connections of area F3 (SMA-proper) and area F6 (pre-SMA) in the macaque monkey. *Journal of Comparative Neurology*. 1993; 338(1): 114–140. [PubMed: 7507940]
- Maeda, S. Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal tract shapes using an articulatory model. In: Hardcastle, WJ.; Marchal, A., editors. *Speech Production and Speech Modeling*. Boston: Kluwer Academic Publishers; 1990.
- Makris N, Kennedy DN, McInerney S, Sorensen AG, Wang R, Caviness VS, et al. Segmentation of subcomponents within the superior longitudinal fascicle in humans: A quantitative, in vivo, DT-MRI study. *Cerebral Cortex*. 2005; 15(6):854–869. [PubMed: 15590909]
- Mathiesen C, Caesar K, Akgoren N, Lauritzen M. Modification of activity-dependent increases of cerebral blood flow by excitatory synaptic activity and spikes in rat cerebellar cortex. *Journal of Physiology*. 1998; 512(Pt 2):555–566. [PubMed: 9763643]
- Matsumoto R, Nair DR, LaPresto E, Bingaman W, Shibasaki H, Luders HO. Functional connectivity in human cortical motor system: a cortico-cortical evoked potential study. *Brain*. 2007; 130(Pt 1): 181–197. [PubMed: 17046857]
- Matsumoto R, Nair DR, LaPresto E, Najm I, Bingaman W, Shibasaki H, et al. Functional connectivity in the human language system: a cortico-cortical evoked potential study. *Brain*. 2004; 127(Pt 10): 2316–2330. [PubMed: 15269116]
- Max L, Guenther FH, Gracco VL, Ghosh SS, Wallace ME. Unstable or insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemporary Issues in Communication Science and Disorders*. 2004; 31:105–122.
- Mazziotta J, Toga A, Evans A, Fox P, Lancaster J, Zilles K, et al. A four-dimensional probabilistic atlas of the human brain. *Journal of the American Medical Informatics Association*. 2001; 8(5): 401–430. [PubMed: 11522763]
- McNeil, MR.; Robin, DA.; Schmidt, RA. *Apraxia of speech: Definition and differential diagnosis*. New York: Thieme; 2007.
- Meyer M, Alter K, Friederici AD, Lohmann G, von Cramon DY. fMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Human Brain Mapping*. 2002; 17(2): 73–88. [PubMed: 12353242]
- Miall RC, Jenkinson EW. Functional imaging of changes in cerebellar activity related to learning during a novel eye-hand tracking task. *Experimental Brain Research*. 2005; 166(2):170–183.
- Miall RC, Wolpert DM. Forward models for physiological motor control. *Neural Networks*. 1996; 9(8):1265–1279. [PubMed: 12662535]
- Middleton FA, Strick PL. Cerebellar output channels. *International Review of Neurobiology*. 1997; 41:61–82. [PubMed: 9378611]
- Mitchell RL, Elliott R, Barry M, Cruttenden A, Woodruff PW. The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia*. 2003; 41(10):1410–1421. [PubMed: 12757912]
- Mochizuki H, Saito H. Mesial frontal lobe syndromes: correlations between neurological deficits and radiological localizations. *Tohoku Journal of Experimental Medicine*. 1990; 161(Suppl):231–239. [PubMed: 2082501]
- Morel A, Kaas JH. Subdivisions and Connections of Auditory-Cortex in Owl Monkeys. *Journal of Comparative Neurology*. 1992; 318(1):27–63. [PubMed: 1583155]
- Mukamel R, Gelbard H, Arieli A, Hasson U, Fried I, Malach R. Coupling between neuronal firing, field potentials, and fMRI in human auditory cortex. *Science*. 2005; 309(5736):951–954. [PubMed: 16081741]
- Nemeth G, Hegedus K, Molnar L. Akinetic mutism associated with bicingular lesions: clinicopathological and functional anatomical correlates. *European Archives of Psychiatry and Neurological Sciences*. 1988; 237(4):218–222. [PubMed: 2849547]

- Neumann K, Preibisch C, Euler HA, von Gudenberg AW, Lanfermann H, Gall V, et al. Cortical plasticity associated with stuttering therapy. *Journal of Fluency Disorders*. 2005; 30(1):23–39. [PubMed: 15769497]
- Nieto-Castanon A, Guenther FH, Perkell JS, Curtin HD. A modeling investigation of articulatory variability and acoustic stability during American English /r/ production. *Journal of the Acoustical Society of America*. 2005; 117(5):3196–3212. [PubMed: 15957787]
- Numminen J, Salmelin R, Hari R. Subject's own speech reduces reactivity of the human auditory cortex. *Neuroscience Letters*. 1999; 265(2):119–122. [PubMed: 10327183]
- O'Reilly JX, Mesulam MM, Nobre AC. The cerebellum predicts the timing of perceptual events. *Journal of Neuroscience*. 2008; 28(9):2252–2260. [PubMed: 18305258]
- Ogawa S, Menon RS, Tank DW, Kim SG, Merkle H, Ellermann JM, et al. Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. A comparison of signal characteristics with a biophysical model. *Biophysical Journal*. 1993; 64(3):803–812. [PubMed: 8386018]
- Ohyama T, Nores WL, Murphy M, Mauk MD. What the cerebellum computes. *Trends in Neurosciences*. 2003; 26(4):222–227. [PubMed: 12689774]
- Ojemann GA. Cortical organization of language. *Journal of Neuroscience*. 1991; 11(8):2281–2287. [PubMed: 1869914]
- Oller DK, Eilers RE. The role of audition in infant babbling. *Child Development*. 1988; 59(2):441–449. [PubMed: 3359864]
- Olthoff A, Baudewig J, Kruse E, Dechent P. Cortical sensorimotor control in vocalization: a functional magnetic resonance imaging study. *Laryngoscope*. 2008; 118(11):2091–2096. [PubMed: 18758379]
- Padoa-Schioppa C, Li CS, Bizzi E. Neuronal activity in the supplementary motor area of monkeys adapting to a new dynamic environment. *Journal of Neurophysiology*. 2004; 91(1):449–473. [PubMed: 12968016]
- Penfield, W.; Rasmussen, T. *The cerebral cortex of man. A clinical study of localization of function.* New York: Macmillan; 1950.
- Penfield, W.; Roberts, L. *Speech and brain mechanisms.* Princeton, N.J: Princeton University Press; 1959.
- Penfield W, Welch K. The supplementary motor area of the cerebral cortex; a clinical and experimental study. *American Medical Association Archives of Neurology and Psychiatry*. 1951; 66(3):289–317.
- Penhune VB, Doyon J. Cerebellum and M1 interaction during early learning of timed motor sequences. *Neuroimage*. 2005; 26(3):801–812. [PubMed: 15955490]
- Perkell JS, Denny M, Lane H, Guenther F, Matthies ML, Tiede M, et al. Effects of masking noise on vowel and sibilant contrasts in normal-hearing speakers and postlingually deafened cochlear implant users. *Journal of the Acoustical Society of America*. 2007; 121(1):505–518. [PubMed: 17297804]
- Perkell JS, Guenther FH, Lane H, Matthies ML, Perrier P, Vick J, et al. A theory of speech motor control and supporting data from speakers with normal hearing and profound hearing loss. *Journal of Phonetics*. 2000; 28:232–272.
- Perkell JS, Guenther FH, Lane H, Matthies ML, Stockmann E, Tiede M, et al. The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *Journal of the Acoustical Society of America*. 2004; 116(4 Pt 1):2338–2344. [PubMed: 15532664]
- Perkell JS, Matthies ML, Tiede M, Lane H, Zandipour M, Marrone N, et al. The distinctness of speakers' /s/-/S/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. *Journal of Speech, Language, and Hearing Research*. 2004; 47(6):1259–1269.
- Pickett E, Kuniholm E, Protopapas A, Friedman J, Lieberman P. Selective speech motor, syntax and cognitive deficits associated with bilateral damage to the putamen and the head of the caudate nucleus: a case study. *Neuropsychologia*. 1998; 36:173–188. [PubMed: 9539237]
- Pihan H, Altenmüller E, Ackermann H. The cortical processing of perceived emotion: a DC-potential study on affective speech prosody. *Neuroreport*. 1997; 8(3):623–627. [PubMed: 9106735]

- Raichle ME, Mintun MA. Brain work and brain imaging. *Annual Review of Neuroscience*. 2006; 29:449–476.
- Reppas JB, Usrey WM, Reid RC. Saccadic eye movements modulate visual responses in the lateral geniculate nucleus. *Neuron*. 2002; 35:961–1074. [PubMed: 12372289]
- Riecker A, Ackermann H, Wildgruber D, Dogil G, Grodd W. Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *Neuroreport*. 2000; 11(9):1997–2000. [PubMed: 10884059]
- Riecker A, Ackermann H, Wildgruber D, Meyer J, Dogil G, Haider H, et al. Articulatory/phonetic sequencing at the level of the anterior perisylvian cortex: a functional magnetic resonance imaging (fmri) study. *Brain and Language*. 2000; 75(2):259–276. [PubMed: 11049668]
- Riecker A, Wildgruber D, Dogil G, Grodd W, Ackermann H. Hemispheric lateralization effects of rhythm implementation during syllable repetitions: An fMRI Study. *Neuroimage*. 2002; 16(1): 169–176. [PubMed: 11969327]
- Rizzolatti G, Arbib MA. Language within our grasp. *Trends in Neurosciences*. 1998; 21(5):188–194. [PubMed: 9610880]
- Rizzolatti G, Camarda R, Fogassi L, Gentilucci M, Luppino G, Matelli M. Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Experimental Brain Research*. 1988; 71(3):491–507.
- Rizzolatti G, Fadiga L, Gallese V, Fogassi L. Premotor cortex and the recognition of motor actions. *Brain Research Cognitive Brain Research*. 1996; 3(2):131–141. [PubMed: 8713554]
- Rizzolatti G, Fogassi L, Gallese V. Parietal cortex: from sight to action. *Current Opinions in Neurobiology*. 1997; 7(4):562–567.
- Robin, DA.; Guenther, FH.; Narayana, S.; Jacks, A.; Tourville, JA.; Ramage, AE., et al. A transcranial magnetic stimulation virtual lesion study of speech; *Proceedings of the Conference on Motor Speech*; Monterey, California. 2008.
- Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature Neuroscience*. 1999; 2(12):1131–1136.
- Ross ED, Mesulam MM. Dominant language functions of the right hemisphere? Prosody and emotional gesturing. *Arch Neurol*. 1979; 36(3):144–148. [PubMed: 435134]
- Roy JE, Cullen KE. Dissociating self-generated from passively applied head motion: neural mechanisms in the vestibular nuclei. *Journal of Neuroscience*. 2004; 24(9):2102–2111. [PubMed: 14999061]
- Schmahmann, J.; Pandya, D. *Fiber pathways of the brain*. New York: Oxford University Press; 2006.
- Schmahmann JD, Pandya DN. The cerebrocerebellar system. *International Review of Neurobiology*. 1997; 41:31–60. [PubMed: 9378595]
- Schreurs BG, McIntosh AR, Bahro M, Herscovitch P, Sunderland T, Molchan SE. Lateralization and behavioral correlation of changes in regional cerebral blood flow with classical conditioning of the human eyeblink response. *Journal of Neurophysiology*. 1997; 77(4):2153–2163. [PubMed: 9114262]
- Shima K, Tanji J. Neuronal activity in the supplementary and presupplementary motor areas for temporal organization of multiple movements. *Journal of Neurophysiology*. 2000; 84(4):2148–2160. [PubMed: 11024102]
- Sidtis JJ, Van Lancker Sidtis D. A neurobehavioral approach to dysprosody. *Seminars in Speech and Language*. 2003; 24(2):93–105. [PubMed: 12709883]
- Simonyan K, Jurgens U. Efferent subcortical projections of the laryngeal motor cortex in the rhesus monkey. *Brain Research*. 2003; 974(1–2):43–59. [PubMed: 12742623]
- Sohn JW, Lee D. Order-dependent modulation of directional signals in the supplementary and presupplementary motor areas. *Journal of Neuroscience*. 2007; 27(50):13655–13666. [PubMed: 18077677]
- Soros P, Sokoloff LG, Bose A, McIntosh AR, Graham SJ, Stuss DT. Clustered functional MRI of overt speech production. *Neuroimage*. 2006; 32(1):376–387. [PubMed: 16631384]

- Sperry RW. Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*. 1950; 43(6):482–489. [PubMed: 14794830]
- Stepniewska I, Preuss TM, Kaas JH. Ipsilateral cortical connections of dorsal and ventral premotor areas in New World owl monkeys. *Journal of Comparative Neurology*. 2006; 495(6):691–708. [PubMed: 16506197]
- Stiller D, Gaschler-Markefski B, Baumgart F, Schindler F, Tempelmann C, Heinze HJ, et al. Lateralized processing of speech prosodies in the temporal cortex: a 3-T functional magnetic resonance imaging study. *Magnetic Resonance Materials in Physics, Biology and Medicine*. 1997; 5(4):275–284.
- Tai YF, Scherfler C, Brooks DJ, Sawamoto N, Castiello U. The human premotor cortex is 'mirror' only for biological actions. *Current Biology*. 2004; 14(2):117–120. [PubMed: 14738732]
- Terband, H.; Maassen, B.; Brumberg, J.; Guenther, FH. A transcranial magnetic stimulation virtual lesion study of speech; Proceedings of the Conference on Motor Speech; Monterey, California. 2008.
- Tesche CD, Karhu JJ. Anticipatory cerebellar responses during somatosensory omission in man. *Human Brain Mapping*. 2000; 9(3):119–142. [PubMed: 10739364]
- Tourville JA, Reilly KJ, Guenther FH. Neural mechanisms underlying auditory feedback control of speech. *Neuroimage*. 2008; 39(3):1429–1443. [PubMed: 18035557]
- Toyomura A, Koyama S, Miyamaoto T, Terao A, Omori T, Murohashi H, et al. Neural correlates of auditory feedback control in human. *Neuroscience*. 2007; 146(2):499–503. [PubMed: 17395381]
- Tseng YW, Diedrichsen J, Krakauer JW, Shadmehr R, Bastian AJ. Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *Journal of Neurophysiology*. 2007; (1):54–62. [PubMed: 17507504]
- Turkeltaub PE, Eden GF, Jones KM, Zeffiro TA. Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *Neuroimage*. 2002; 16(3 Pt 1):765–780. [PubMed: 12169260]
- Urban PP, Marx J, Hunsche S, Gawehn J, Vucurevic G, Wicht S, et al. Cerebellar speech representation: lesion topography in dysarthria as derived from cerebellar ischemia and functional magnetic resonance imaging. *Archives of Neurology*. 2003; 60(7):965–972. [PubMed: 12873853]
- Van Buren JM. Confusion and disturbance of speech from stimulation in the vicinity of the head of the caudate nucleus. *Journal of Neurosurgery*. 1963; 20:148–157. [PubMed: 14192084]
- Villacorta VM, Perkell JS, Guenther FH. Sensorimotor adaptation to feedback perturbations on vowel acoustics and its relation to perception. *Journal of the Acoustical Society of America*. 2007; 122(4):2306–2319. [PubMed: 17902866]
- Viswanathan A, Freeman RD. Neurometabolic coupling in cerebral cortex reflects synaptic more than spiking activity. *Nature Neuroscience*. 2007; 10(10):1308–1312.
- von Holst E, Mittelstaedt H. Das reafferenzprinzip. *Naturewissenschaften*. 1950; 37:464–476.
- Voss M, Ingram JN, Haggard P, Wolpert DM. Sensorimotor attenuation by central motor command signals in the absence of movement. *Nature Neuroscience*. 2006; 9(1):26–27.
- Walker JP, Pelletier R, Reif L. The production of linguistic prosodic structures in subjects with right hemisphere damage. *Clinical Linguistics and Phonetics*. 2004; 18(2):85–106. [PubMed: 15086132]
- Wambaugh JL, Duffy JR, McNeil MR, Robin DA, Rogers MA. Treatment guidelines for acquired apraxia of speech: Treatment descriptions and recommendations. *Journal of Medical Speech-Language Pathology*. 2006; 14(2):Xxxv–Lxvii.
- Wildgruber D, Ackermann H, Grodd W. Differential contributions of motor cortex, basal ganglia, and cerebellum to speech motor control: effects of syllable repetition rate evaluated by fMRI. *Neuroimage*. 2001; 13(1):101–109. [PubMed: 11133313]
- Williamson JB, Harrison DW, Shenal BV, Rhodes R, Demaree HA. Quantitative EEG diagnostic confirmation of expressive aprosodia. *Appl Neuropsychol*. 2003; 10(3):176–181. [PubMed: 12890643]

- Wise RJ, Greene J, Buchel C, Scott SK. Brain regions involved in articulation. *Lancet*. 1999; 353(9158):1057–1061. [PubMed: 10199354]
- Wolpert D, Miall C, Kawato M. Internal models in the cerebellum. *Trends in Cognitive Science*. 1998; 2(9):338–347.
- Wolpert DM, Kawato M. Multiple paired forward and inverse models for motor control. *Neural Networks*. 1998; 11(7–8):1317–1329. [PubMed: 12662752]
- Ziegler W, Kilian B, Deger K. The role of the left mesial frontal cortex in fluent speech: evidence from a case of left supplementary motor area hemorrhage. *Neuropsychologia*. 1997; 35(9):1197–1208. [PubMed: 9364490]

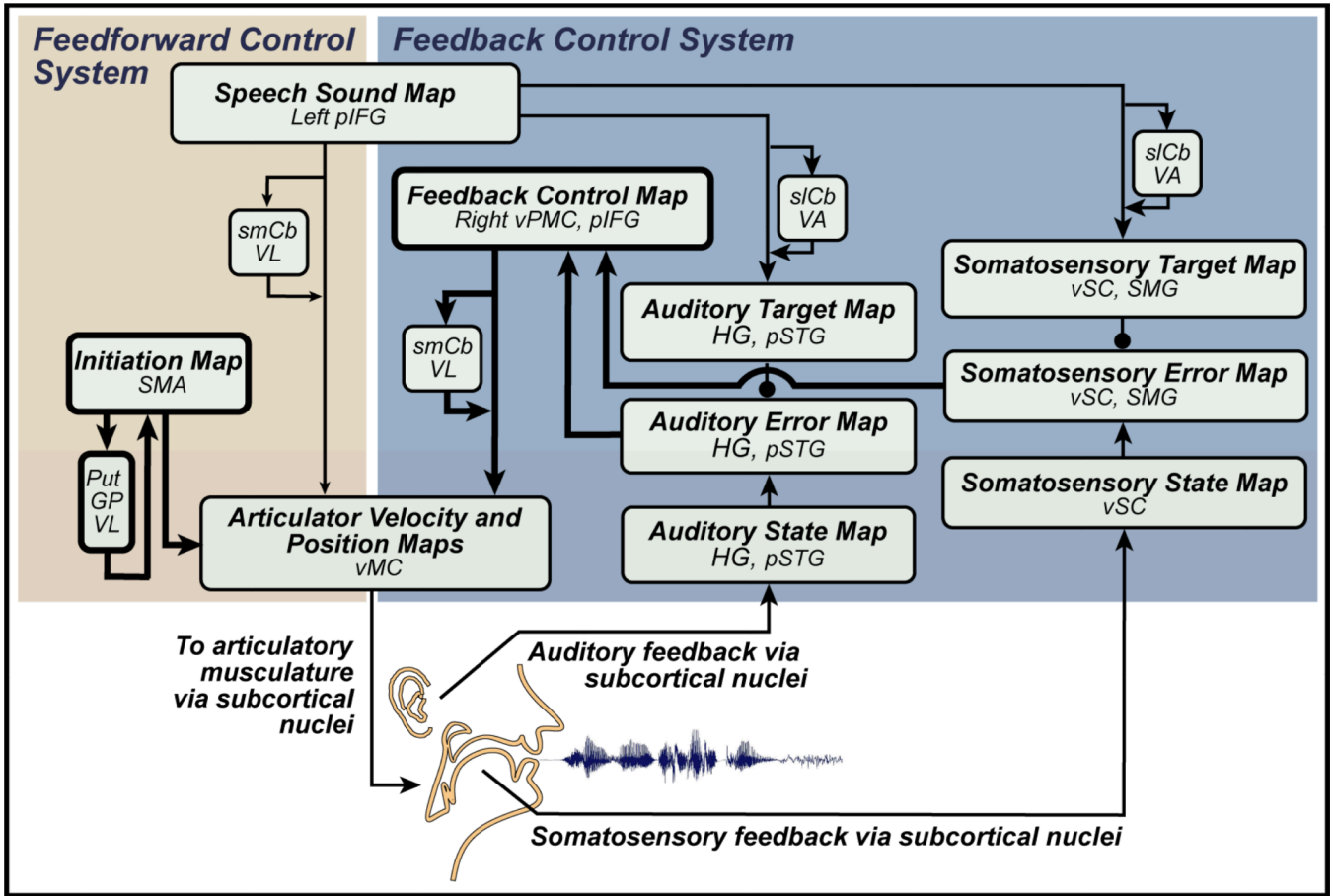


Figure 1. The DIVA model of speech acquisition and production. Recently added modules and connections are highlighted by black outlines. Model components associated with hypothesized neuroanatomical substrates. *Abbreviations:* GP = globus pallidus; HG = Heschl's gyrus; pIFg = posterior inferior frontal gyrus; pSTg = posterior superior temporal gyrus; Put = putamen; sICB = superior lateral cerebellum; smCB = superior medial cerebellum; SMA = supplementary motor area; SMG = supramarginal gyrus; VA = ventral anterior nucleus of the thalamus; vMC = ventral motor cortex; vPMC = ventral premotor cortex; vSC = ventral somatosensory cortex.

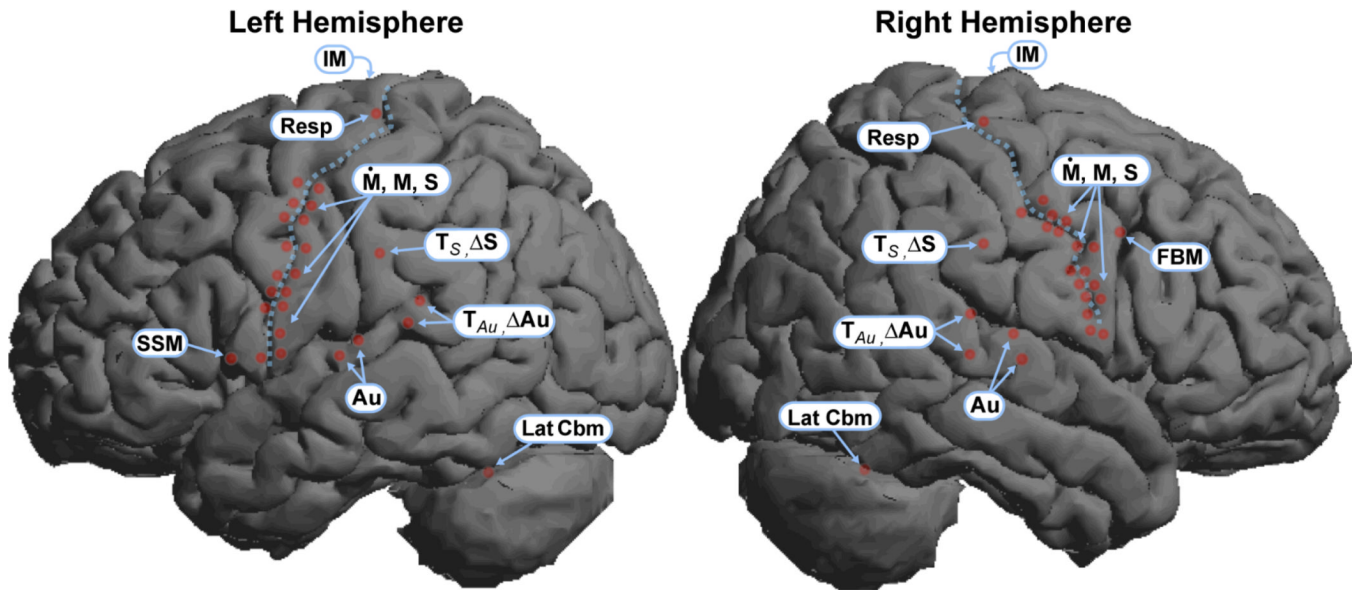


Figure 2.

Neuroanatomical mapping of the DIVA model. The location of DIVA model component sites (red dots) are plotted on renderings of the left (top) and right (bottom) lateral surfaces of the SPM2 canonical brain. Sites immediately anterior to the central sulcus (dotted line) represent cells of the model's articulator velocity (\dot{M}) and position (M) maps. Sites located immediately posterior to the central sulcus represent cells of the somatosensory state map (S). Subcortical sites (basal ganglia, thalamus, paravermal cerebellum, deep cerebellar nuclei), are not shown. *Additional abbreviations:* Au = auditory state map; ΔAu = auditory error map; FB = feedback control map; IM = initiation map; Lax_{int} , Lax_{ext} = intrinsic and extrinsic larynx; Lat Cbm = lateral cerebellum; Resp: respiratory motor cells; ΔS = somatosensory error map; SSM = speech sound map; T_{Au} = auditory target map; T_S = somatosensory target map.

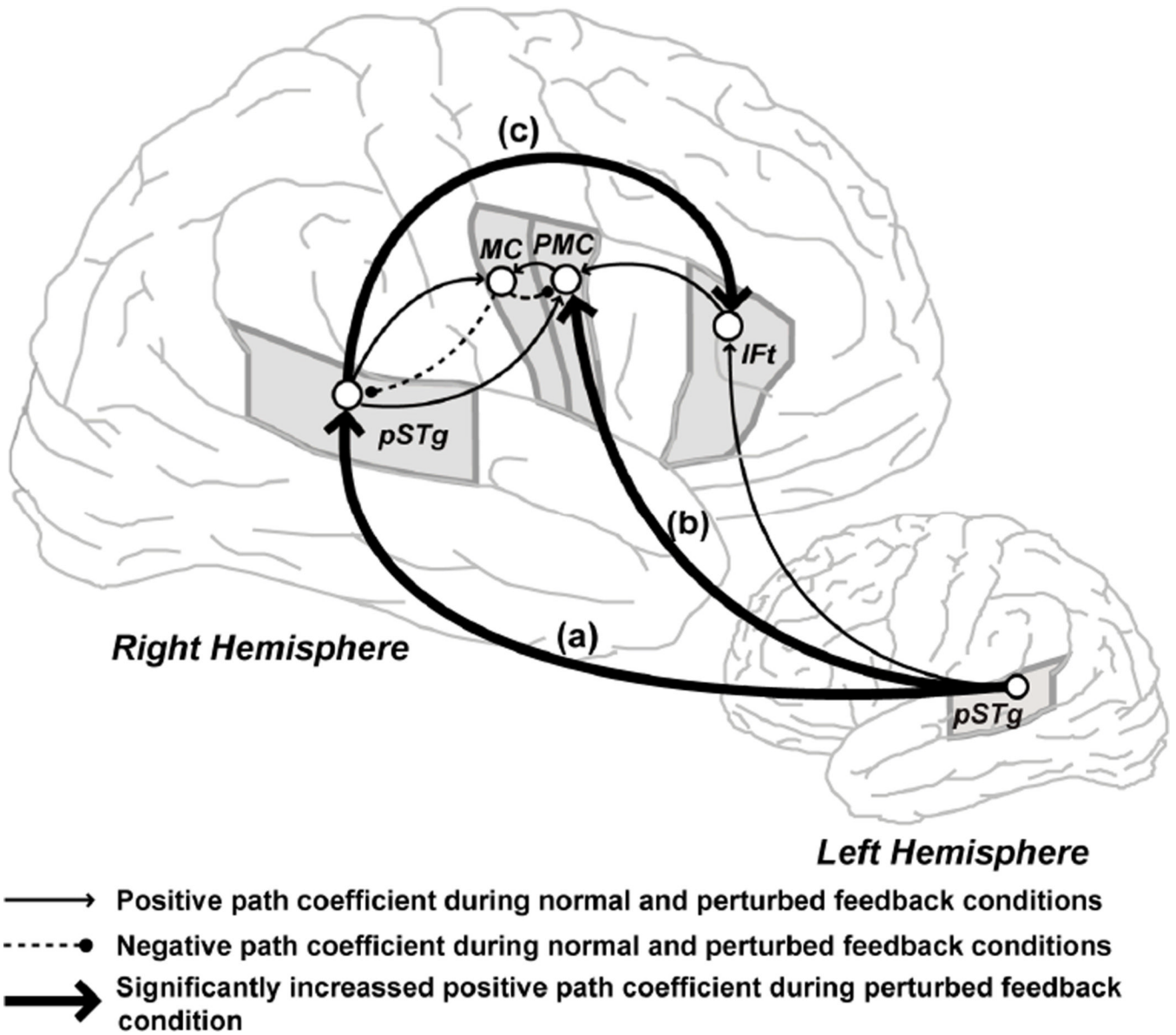


Figure 3. Effective connectivity within the auditory feedback control network. Structural equation modeling demonstrated significant modulation of interregional interactions within the schematized network when auditory feedback was perturbed during speech production. Pairwise comparisons of path coefficients in the normal and perturbed feedback conditions revealed significant increases in the positive weights from left posterior superior temporal gyrus (pSTg) to right pSTg (the path labeled a in the diagram above), from left pSTg to right ventral premotor cortex (PMC; path b), and from right pSTg to right inferior frontal gyrus, pars triangularis (path c) when auditory feedback was perturbed during speech production. *Additional abbreviation:* MC = motor cortex.

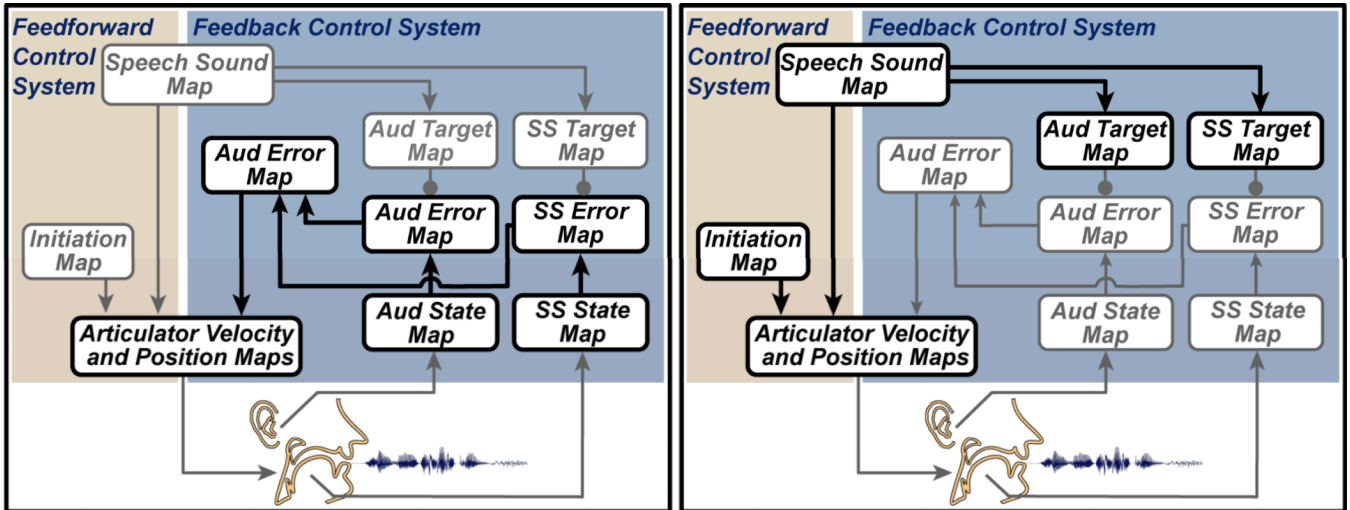


Figure 4. Learning in the DIVA model. Simplified DIVA model block diagrams indicate the mappings that are tuned during the two learning phases (heavy black outlines). *Left:* Early babbling learning phase. Pseudo-random motor commands to the articulators are associated with auditory and somatosensory feedback. The paired motor and sensory signals are used to tune synaptic projections from sensory error maps to the feedback control map. The tuned projections are then able to transform sensory error inputs into feedback-based motor commands. *Right:* Imitation learning phase. Auditory speech sound targets (encoded in projections from the speech sound map to the auditory target map) are initially tuned based on sample speech sounds from other speakers. These targets, somatosensory targets, and projections in the feedforward control system are tuned during attempts to imitate a learned speech sound target.

Table 1

The location of new DIVA model components in MNI space.

Model Components	Left Hemisphere			Right Hemisphere		
	x	y	z	x	y	z
<i>Feedback Control Map</i>						
Right ventral premotor cortex				60	14	34
<i>Initiation Map</i>						
SMA	0	0	68	2	4	62
Putamen	-26	-2	4	30	-14	4
Globus Pallidus	-24	-2	-4	24	2	-2
Thalamus	-10	-14	8	10	-14	8
<i>Articulator Velocity and Position Maps</i>						
Larynx (Intrinsic)	-53	0	42	53	4	42
Larynx (Extrinsic)	-58.1	6.0	6.4	65.4	5.2	10.4
<i>Somatosensory State Map</i>						
Larynx (Intrinsic)	-53	-8	42	53	-14	38
Larynx (Extrinsic)	-61.8	1	7.5	65.4	1.2	12